

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2002, 02:56:10 ; Search time 1665 Seconds

(without alignments)  
16541.650 Million cell updates/sec

Title: US-09-716-536-7

Perfect score: 2007

Sequence: 1 gtgcgtgagcgaatttc.....aaaaaaaaaaaaaaaaaa 2007

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_estlum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estcpl:\*  
7: em\_estcro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1075.2	53.6	2696	11	AK012948 Mus muscu
2	1067.2	53.2	1938	11	AK012786 Mus muscu
3	894.4	44.6	959	9	AL560947 AL560947
4	863.4	43.0	887	9	AL560912 AL560912
5	777.7	38.7	1061	13	BM464844 BM464844
6	685.8	34.2	775	9	AL582241 AL582241

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	MEDLINE	PUBMED
AK012948	2696 bp	musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810054N23;TRAF-interacting protein, full insert sequence.	AK012948.1	GI:12850018	HTC; CAP trapper.	Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library	Mus musculus	1	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)	JOURNAL MEDLINE	20493774	11042159
AK012948	2696 bp	musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810054N23;TRAF-interacting protein, full insert sequence.	AK012948.1	GI:12850018	HTC; CAP trapper.	Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library	Mus musculus	1	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)	JOURNAL MEDLINE	20493774	11042159

## ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	MEDLINE	PUBMED
AK012948	2696 bp	musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810054N23;TRAF-interacting protein, full insert sequence.	AK012948.1	GI:12850018	HTC; CAP trapper.	Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library	Mus musculus	1	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)	JOURNAL MEDLINE	20493774	11042159



```

QY 418 ACTTGCAGGATACGCTGGAGAAGAACGCAATGCTACTGTGTATCTCTGCAGCAGCCTTG 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 427 ACTCTACGGGACACCCCTGGAGAAGACCAATGCTACCGTGGAGTCCCTACAGAAAGCCTTA 486
QY 478 GCGAAGCCGAGATGCTGTGCTCCACACTGAAAGAACAGATGAATGACTTACAGCAGCAG 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 487 AACCAAGGCGAGATGCTGTGCTCCACCCCTGAAAAAACAGATGAATGCTGTGAGCAGCG 546
QY 538 CAGGATGACACCAACACACAGAGAGGAGGCGCGGCTCAGAGGCAAGATGAAGACC 597
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 547 CAGGATGACACCAACACACCTGGAGAGGAGGCCACCGACTCAGAGTGAATGAAGAACCC 606
QY 598 ATGAGCAGATTTAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGGAGAGATGATCCGA 657
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 607 ATGAGACCAATTTAGCTTCTACTCCAGAGCCAGCGGCTTGAGGTGGAGAGATGATCCGA 666
QY 658 GACATGGCTGTGGAGACGTCAGCGGTGGAAACAGCTGCTGTGATCTGTCTCTCAG 717
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 667 GACATGGGTGTGGAGACGTCAGCGGTGGAGACAGCTGCTGTGATCTGTCTCTCAG 726
QY 718 AAGAGTACGAGATCTAAAGAGGAGGAGGAGGCTCAGGGGAGGTGGCTGACAGCTG 777
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 727 AAGAGTATGAGATCTGAAGGAGCTCGAAGGCCACAGGGGAACTGGCTGACAGCTTG 786
QY 778 AGGAAGATTTTCTTCTCCAGAAAGCAATTGACAGACAGCTTCTGTAATTTGATCAG 837
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 787 AAGAAAGATTTGCTCTCTAGAGCAAGTTGAAGACTCTCAACACTAGAGCTGAGTCA 846
QY 838 GCCAATTTGAATGAACTGACAGCCCAAGAAAGACTTACAGAGTCTGACAAAGAAATCATG 897
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 847 GCCAATTTGAATGAGTGAAGCCCAAGAAAGACTTACAAAGTCTGACCAAGAGATCAG 906
QY 898 AGCCTAAAAAGAGTAAGTGAATGTCAGAGAAACCTTGACCTGACACAGAGGCGAGT 957
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 907 AGCCTAAGAAAGAGTGAATGATCTCTCAGAGAACTTGAAGCTGCTCGGCGACCAAT 966
QY 958 GAGACTGTGACCGCTGCTTTTAAAGAGCCCAAGCCCTGTGGA---GGTAAATTCGAAG 1014
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 967 GAGAGGCTAGCGCGCTGCTTTTAAAGAGCCCAAGCCCTGTGGAATGATGAAGCCGAGG 1026
QY 1015 CTCGCCGCGCATCTCTCGGTGATGATTAATGATCTCAATGCTCACTTTGATGTGATCT 1074
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1027 CTCACACACCAACCCCTCGGTGATGATGATCAATACCACTTTGATGTAATTAACC 1086
QY 1075 CCCCCAGCCGCGCTTCAGCTCCAGCATGGTTACTAGCAAAAACCTTGCCCTAGGAAG 1134
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1087 CCTCCAAACCCAGACTCTGTGCTCCAGCATTTGCTCCCAAGAGCTGTGCTGGAGAGG 1146
QY 1135 TCACACTCCCAATTCAGATGTGCCAGAAAGATATGCAAAAGCCCAAGAGAGAGTCC 1194
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1147 GCACGCTTCACATGAGATGTCTCAAGAAAGGTGCACAAATCTCCAAAGCCGAGTCC 1206
QY 1195 CAGCTCTCACTGGGTGGCCAGAGCTGTGACAGAGAGCCAGATGAGCAACTGTTGGTCC 1254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1207 CAGCTCTCACTGGGTGGCCAGAGCTGTGAGAGACTAGATGAGGAACCTGTGCTGCC 1266
QY 1255 TTCCTAATTTTGTCCGAATGCATCTAGGCGAAGAACGCCCAAGGCCAGGTCA 1314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1267 TTCCTCTTTCACTCCGGAATGCTGTCTGGGCAAAACAGCCCAAGAGACACACAGA 1326
QY 1315 GAGTCTCTTTCAGCAAAAGATGTGTAAAGACAGCCTTGATGGGTGCGTCCGCGAGCA 1374
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1327 GATCCCGTTCAGACACAGATGTGTAAAGATGAGCTTTGATGGGTGTGAGAGACAGACA 1386
QY 1375 AATTTATCCAGCTTCTGACACAGTATGATCCGCCATTTGCTTTAAGCCCAAGACC 1434
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1387 AATTTATCCAGCTTCTGACACACATTTATCCAGCAGCTGCTTTAAGTCCAAAGGCC 1446
QY 1435 AAGGTTAAGCAGAGGTTGAAGAGCGTTCCTCTCTCTTCAGGCGCAAGCTGGAC 1494
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1447 AAGAGTAAACAGAAATGAGAAATTAAGACTGTGATTTGCTCCAGGCCCAAGCTGGAT 1506

```

```

QY 1495 ACCTTCTGTGTCGTGAGAACAGTGAATCTGACCAATGGCCAGACATGCTTCGCACT 1554
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1507 ACCTTCTGTATGTCAG-----TGAGCGTACACAGATCATGTTTGCATTT 1551
QY 1555 TGTAGTCAAGACACTGTCCAGCAGGCTTTGTGTACAGAGCCCTTACTTTGGAGCAGCC 1614
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1552 AGTGGGCCAAAGACTGTGGCTAACCGGAAGTGTTTTGGAAAGATGGCTCTCTTG----- 1604
QY 1615 TGAGTGTAAAGGCAACAAACAGGTGAGGTGATGATACACCCAGACAGTCTCTCC 1674
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1605 -----GACCACTCCAAAGAGATGCGCCAGAAACACACTTCC 1641
QY 1675 TGCCCTACCCCTGCCCACTCTACGACTGGAGAGTGAATGACACGACCCACTGATCTG 1734
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1642 TGTGTACTGTGGCGCTTGACAC--ACATGGGAAGCCACATGACAGTTTACTGTTCGA 1699
QY 1735 TCAGCAGTCTCTGC-TCTGTGCGCAGGCTCTTTATAGCCATGATCAAGATGTGTCAG 1793
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1700 TCGGCAGGCGCTACTTCCAGTTGACAGGTTTGTCTTATAGTACAAACAGGTGTGCTGG 1759
QY 1794 ACTCTTCTGGGCGCTGAGAGCACCGGTCACTT 1825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1760 ACTCTTTTGTTTTATAGAACAGGCTCAAT 1791

```

```

RESULT 2
AK012786
LOCUS
DEFINITION
Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2810021M06:TRAF-interacting protein, full
insert sequence.
ACCESSION
AK012786
VERSION
AK012786.1 GI:12849758
KEYWORDS
HTC: CAP trapper.
Mus musculus (Strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA,
clone:lib:RIKEN full-length enriched mouse cDNA library
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ORGANISM
Mus musculus
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL
99279253
MEDLINE
10349636
PUBMED
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
20499374
MEDLINE
11042159
PUBMED
REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer
Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL
20530913
MEDLINE
11076861
PUBMED
REFERENCE
4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, T.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,

```



OY	806	AGTTGCACGAGCTTCTACTGTGAATGGATTCAGGCCAAGTTAGCACTGAAGTCCAGCCAGA	865
Db	798	AGTTGAAGACTCTTCACACTGAGCTGAGTCCAGGCCAAGTTAACTGAGGTGAGGCCAGA	857
OY	866	AGGACTTACAGAGTGCTGACACAGAAATCATGAGCCTGAAAAAGAAAGCTAACGATGCTGC	925
Db	858	AGSACTTACAAGAGTGCGACACAGAGATCACGAGAGCCTTAAGAAAAGAAAGCTGATACCTGCC	917
OY	986	GCCCAGCCCCCTGTGGA---GGTAGATCTGAAGCTCCGCGCGGAGCATCTCTCCGTATGTATA	1042
Db	978	GCCCAGCCCCCTGTGTGAGATGATATGACCAGAGCTTCCAGGCACTCTGCGGTATGTAGA	1037
OY	1043	TTCGATCTCAATGCTACTCTTGTATGTGATACTGCCACGCCGCGGCTCCAGCTCCAGC	1102
Db	1038	TTCGATCTCAATACCACTTGTGATGTAAATACCCCTTCCAACCCAGACCTCTGCTCCAGC	1097
OY	1103	ATTGTTACTACGAAAAAAGCTTGGCTGTGAGAAAGCACACACACCCCAATTCAGATGATGCCCA	1162
Db	1098	ATTGCTCTCCCAAGAAAGCTGTGCTGTGAGAGGAGCACAGCTCTCCATGAGAGATGTCTCA	1157
OY	1163	AGAGATATGCAAAAGGCCCCAGAGAAAGAGTCCAGCTCACTGTGGTGGCAGAGACTGTG	1222
Db	1158	AGAGATGTCACAAAGTCTCCAAAGCCGAGTCCAGCTCACTGTGGTGGCAGAGAGTGTG	1217
OY	1223	CAGAGAGCCAGATGAGAGAACTGGTGTGTGCTCTCCCTATTTTTTCTCCGAATGCATCC	1282
Db	1218	TAGGAGAGCTAGATGAGAGAACTGGCTGTGTGCTCTCTCTCTTCTTCCGGAATGTGTCC	1277
OY	1283	TAGGCCAGAAACAGCCCAAAAGSCCCAGGCTCAGAGTCTCTTGACGACAAAGTGGTAA	1342
Db	1278	TGGTTCAGAAACAGCCCAAGAGGACCAAGACAGAAATCCCGTTGCAGACAGATGTGGTAA	1337
OY	1343	GGACAGCCTTCATGGGCTCGGTGGGCCCGGACAAATTCATCCAGCCTTACTGCACACATCA	1402
Db	1338	GAAATAGCCTTGTATGGGCTTGGAGGACGAAACAAATTCATCCAGCCTTGGGACACAAACA	1397
OY	1403	TGATGTCGCCCATTTGGCCGTGAATACCCCAAGGCTTAGAGAGAGGTTAGAGGGTGAAGA	1462
Db	1398	TTATTCACCAAGATGCTGTTAAAGTCCAGAGCCAAAGATGAACAGAAAGTGAATAAAGA	1457
OY	1463	CCGTGCTTCTCTCTCCAGGCCCAAGCTGGACACCTTCTGTGTGCTGAGAAACAGTAG	1522
Db	1458	CTGTGATGTTCTGCTCTCCACGCCCAACCTGGATACCTTTTATGTGAG-----	1504
OY	1523	TCTGACCAATGCGCAAGCANATGCTGTCGAACCTGTATGTCAAGGACTGTCCAGGCAAGGT	1582
Db	1505	--TGAACGGTGACCAAGAGTCAATTTTTCGAATTAAGTGGCAAGACTGTGCTAAACCGTAG	1562
OY	1583	TTTGTGACAGAGCCTTACTTTCGGAGCAAGCCTGAGAGGTGAAGGCAACAACAGGTGA	1642
Db	1563	TGTTTTTGAAGAATGGGTCTCTTG-----GACCA	1592
OY	1643	GGGTGATGTGACACCCAGAGACTGCTTTCCTGCCCCACCTCGCCCACTCTTAAGAC	1702
Db	1593	GTTCCAAAGAGATGGCCAGAAAACACACTTCTGTGTTCACGTCCGCTCGAGAC--ACAC	1650
OY	1703	TGGAGGTGACATGACAGCCCACTGATCTCTGACGAGAGGTCTGCG--TCTGTGGCAAGC	1761
Db	1651	TGGGAAGCCACATGACCAAGTTTACTGTTCGATACAGAGGCTTACTTTCACATTGACAGGG	1710
OY	1762	TCTTGTATATGCAATGATCAGATGTGTGTCAGACTCTTCTGTGGGCTGTGAGACCAAGCTC	1821
Db	1711	TTTTGCTTATATGACAAACAGAGTGTGGCTGAGACCTCTTTGTTTTTAAAGAACAGGCTC	1770
OY	1822	ACTT 1825	
Db	1771	ACAT 1774	

LOCUS	AL560947	959 bp	mRNA	linear	EST 16-FEB-2001
DEFINITION	AL560947 LTI_NFL010_BC2	Homo sapiens cDNA	clone CS0DL005YM09 5		
ACCESSION	AL560947				
VERSION	AL560947.1	GI:12907896			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 959)				
TITLE	Li,W.B., Gruber,C., Jeesee,J. and Polayes,D.				
JOURNAL	Full-length cDNA libraries and normalization unpublished (2001)				
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: <a href="mailto:segre@genoscope.cns.fr">segre@genoscope.cns.fr</a> , <a href="http://www.genoscope.cns.fr">www.genoscope.cns.fr</a> .				
FEATURES					
Source	1..959 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="CS0DL005YM09" /clone_lib="LTI_NFL010_BC2" /sex="male" /issue_type="B cells from Burkitt lymphoma" /note="Vector: pCMVSPORT 6; Site:1: NotI; 1st strand cDNA was primed with a NotI-Oligo(gt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : <a href="mailto:fliang@lifetech.com">fliang@lifetech.com</a> URL : <a href="http://fulllength.invitrogen.com">http://fulllength.invitrogen.com</a> "				
BASE COUNT	259 a 237 c 275 g 185 t	3 others			
ORIGIN					
Query Match	44.6%; Score 894.4; DB 9; Length 959;				
Best Local Similarity	97.1%; Pred. No. 1.2e-152;				
Matches	927; Conservative	3; Mismatches	23; Indels	2; Gaps	2;
21	AAGCAAGCGGAGGCGGGCGCTCTACGAAGCGGACCTGTACACTTTCTTTGGCTGCT	80			
DB	6 AAGCGAGGCGCGGGCGCGCTCTACGAAGCGGACCTGTACACTTTCTTTGGCTGCT	65			
81	GGGCGCTTGAATCCAGGCATCATCTCGTCTGTGCCTATCTGCTCGACTT	140			
DB	66 GGGCGCTTGAATCCAGGCATCATCTCGTCTGTGCCTATCTGCTCGACTT	125			
141	CTTGCATCACTCCCGAGCTGGCGGCATTCACATGCGGCCACACCTTCCACTTGCAGTG	200			
DB	126 CTTGCATCACTCCCGAGCTGGCGGCATTCACATGCGGCCACACCTTCCACTTGCAGTG	185			
201	CCTAATTCAGTCTTTGAGACAGACCAAGTGGACCTGCGCCAGTCCGCAATCCAGT	260			
DB	186 CCTAATTCAGTCTTTGAGACAGACCAAGTGGACCTGCGCCAGTCCGCAATCCAGT	245			
261	TGGCAAGAAGCAATTAATCAATTAACCTCTTTTATCTTGGCCAGAGAGAGAAATGT	320			
DB	246 TGGCAAGAAGCAATTAATCAATTAACCTCTTTTATCTTGGCCAGAGAGAGAAATGT	305			
321	CTTGCATCGAGATTTCTTAAGATGAATGTGACATGTACAGGCCAGCTTTCCAGAA	380			
DB	306 CTTGCATCGAGATTTCTTAAGATGAATGTGACATGTACAGGCCAGCTTTCCAGAA	365			
381	AGACAAGAGAAAGAGACAGCGAGCTCATATGCAACTGTGGCGGATTAAGCTTGAGAA	440			
DB	366 AGACAAGAGAAAGAGACAGCGAGCTCATATGCAACTGTGGCGGATTAAGCTTGAGAA	425			



```

DEFINITION  AGENCOURT_6428661 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5504946
5', mRNA sequence.
ACCESSION   BM464844
VERSION     BM464844.1 GI:18513886
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
NIH-MGC http://mhc.nci.nih.gov/.
Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT     Email: Robert.Strausberg, Ph.D.
            Email: c990bs@remail.nih.gov
            Tissue Procurement: Lou Staudt
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Place: LBL/ML2147 row: b column: 19
            High quality sequence start: 2
            High quality sequence stop: 641.
            Location/Qualifiers
                1..1061
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5504946"
                /clone_lib="NIH_MGC_85"
                /tissue_type="Lymphoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: lymph; Vector: pCMV-Sport6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                Average insert size 1.867 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC Library."
BASE COUNT  282 a 272 c 297 g 202 t 8 others
ORIGIN
Query Match      38.7%; Score 777; DB 13; Length 1061;
Best Local Similarity 97.2%; Pred. No. 2.3e-131;
Matches 808; Conservative 0; Mismatches 21; Indels 2; Gaps 2;

QY 34 CGGGGGCTCTAGCAAGCCGAGCTTCTAGCAATTTCTTGGCTGCTGGCCCTTGAGT 93
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 26 CGGGCGCGCTCTAGCAAGCCGAGCTTCTAGCAATTTCTTGGCTGCTGGCCCTTGAGT 85
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 94 CCAGCATCATGCTTATCGTGTCTGTGTGCAATCTGCTCGAGCTTCTTGATCACTGC 153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 86 CCAGCATCATGCTTATCGTGTCTGTGTGCAATCTGCTCGAGCTTCTTGATCACTGC 145
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 154 CGGAGGTGGCCGCAATCAGTGGGCGACACCTTCCACTTGGAGTGGCTTAATTCAGTCC 213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 146 CGGAGGTGGCCGCAATCAGTGGGCGACACCTTCCACTTGGAGTGGCTTAATTCAGTGG 205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 214 TTGGAACAGCAGCAAGTGGAGCTTGGCCAGATGCCGAATTCAGTGGCAAAAGAAC 273
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 206 TTGGAACAGCAGCAAGTGGAGCTTGGCCAGATGCCGAATTCAGTGGCAAAAGAAC 265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 274 ATTATGAATTAAGCTTCTTGTGATCTTGGCCAGAGAGAGAAATGCTGATCGAGAA 333
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 266 ATTATGAATTAAGCTTCTTGTGATCTTGGCCAGAGAGAGAAATGCTGATCGAGAA 325
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 326 TTCTTAAGAATGAATGAGCAATGTCAGAGCCAGCTTCCAGAAAGAGAGAGAGAA 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 334 TTCTTAAGAATGAATGAGCAATGTCAGAGCCAGCTTCCAGAAAGAGAGAGAGAA 393
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 326 TTCTTAAGAATGAATGAGCAATGTCAGAGCCAGCTTCCAGAAAGAGAGAGAGAA 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 394 CGAGACAGCAGCTCATCTGCACTTGGCGGATRACGCTGGAAGAGCAATGCTACT 453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 386 CGAGACAGCAGCTCATCTGCACTTGGCGGATRACGCTGGAAGAGCAATGCTACT 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 454 GTGCTATCTCTGAGCAGGCGCTTGGCAAGCCGAGATGCTGTCCACTGAAGAAAG 513
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 446 GTGCTATCTCTGAGCAGGCGCTTGGCAAGCCGAGATGCTGTCCACTGAAGAAAG 505
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 514 CAGATGAAGTACTTAAAGCAGCAGCAGATGAGACCAACCAAGCAGAGAGCGGCG 573
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 506 CAGATGAAGTACTTAAAGCAGCAGCAGATGAGACCAACCAAGCAGAGAGCGGCG 565
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 574 CGGCTCAGCAGCAAGTGAAGACCATGAGCAGATTTGAGCTTTACTCCAGAGCCAGCTC 633
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 566 CGGCTCAGCAGCAAGTGAAGACCATGAGCAGATTTGAGCTTTACTCCAGAGCCAGCTC 625
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 634 CCGAGCTGAGAGAGATGATCCGAGACATGGGCTGGAGACAGCAGCGGGAACAGCTG 693
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 626 CCGAGCTGAGAGAGATGATCCGAGACATGGGCTGGAGACAGCAGCGGGAACAGCTG 685
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 694 GCTGTACTGTGTCTCTCAAGAAAGATGAGCAAGTCTAAAGAGCAGCAGGAGCC 753
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 686 GCTGTACTGTGTCTCTCAAGAAAGATGAGCAAGTCTAAAGAGCAGCAGGAGCC 745
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 754 TCAAGGAGAGTGGCTACAGCTGAGAGAGATTTGTTCC-TCCAGAAAGCAAGTTGCA 812
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 746 TCAGGGAGAGTGGCTACAGCTGAGAGAGATTTGTTCC-TCCAGAAAGCAAGTTGCA 805
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 813 GACAGTCTACTCTGAATTTGATCAGGCAAG-TTGAAGTGAAGTCAAGCC 862
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 806 GACAGTCTACTCTGAATTTGATCAGGCAAG-TTGAAGTGAAGTCAAGCC 856
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
AL582241/C 775 bp mRNA linear EST 16-FEB-2001
LOCUS      AL582241.L1L_NFL010_BC2 Homo sapiens cDNA clone CS0DL005Y108 3
DEFINITION AL582241.L1L_NFL010_BC2 Homo sapiens cDNA clone CS0DL005Y108 3
ACCESSION  AL582241
VERSION     AL582241.1 GI:12950030
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 775)
Li W.B., Gruber C., Jesse J., and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..775
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DL005Y108"
/clone_lib="L1L_NFL010_BC2"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT  156 a 199 c 217 g 192 t 11 others
ORIGIN
Query Match      34.2%; Score 685.8; DB 9; Length 775;
Best Local Similarity 95.2%; Pred. No. 9e-115;
Matches 738; Conservative 11; Mismatches 18; Indels 8; Gaps 4;

```



```

Oy 1045 GATTCATATGTAACCTTTGATGTGATTAATCCCGAGCCCGCCCTCCAGCTCCAGCAT 1104
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 775 GATTCATATGTAACCTTTGATGTGATTAATCCCGAGCCCGCCCTCCAGCTCCAGCAT 716
Oy 1105 GGTACTAGAAAAAATTTGGCTAGAGAAAGTCACTATCCCAATTAGAGATGTCCTCAAG 1164
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 715 GGTACTAGAAAAAATTTGGCTAGAGAAAGTCACTATCCCAATTAGAGATGTCCTCAAG 656
Oy 1165 AAGATATGCAAGAGCCCGAGAGAGATGCCAGCTCTCACTGGGTGGCCAGAGCTGTGA 1224
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 655 AAGATATGCAAGAGCCCGAGAGAGATGCCAGCTCTCACTGGGTGGCCAGAGCTGTGA 596
Oy 1225 GGAGAGCCAGATGAGAACTGTTGGTGCCTTCCCTATTTTGTCCGGAATGCCATCTTA 1284
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 595 GGAGAGCCAGATGAGAACTGTTGGTGCCTTCCCTATTTTGTCCGGAATGCCATCTTA 536
Oy 1285 GGCAGAAACAGCCCAAGAGCCCGAGTCAAGTCTCTTGCAGCAAAAGATGTGTAAAG 1344
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 535 GGCAGAAACAGCCCAAGAGCCCGAGTCAAGTCTCTTGCAGCAAAAGATGTGTAAAG 476
Oy 1345 ACAGGCTTCATGGGCTGGTGGCGGAGCAAAATTCATCAGGCTTACACAGTCATG 1404
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 475 ACAGGCTTCATGGGCTGGTGGCGGAGCAAAATTCATCAGGCTTACACAGTCATG 416
Oy 1405 ATCCGCCCATTTGCTGTTAAAGCCCAAGACAGTTAAAGAGGTTGAGGTTGAAGACC 1464
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 415 ATCCGCCCATTTGCTGTTAAAGCCCAAGACAGTTAAAGAGGTTGAGGTTGAAGACC 356
Oy 1465 GTGCTTCTCTCTTCCAGGCGCAAGCTGACACCTTCTGTGTGAGAACAGTGAATC 1524
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 355 GTGCTTCTCTCTTCCAGGCGCAAGCTGACACCTTCTGTGTGAGAACAGTGAATC 296
Oy 1525 TCACCAATGGCCAGACATGCTGCACTTGTAGTGAAGGACGTGCCAGGAGGG--T 1582
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 295 TCACCAATGGCCAGACATGCTGCACTTGTAGTGAAGGACGTGCCAGGAGGGTT 236
Oy 1583 TTGTGACAGAGCCCTTACTTTGGGACAGCCTGAGGTGAAGGGAGCAAAACAGTGA 1642
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 235 TTGTGACAGAGCCCTTACTTTGGGACAGCCTGAGGTGAAGGGAGCAAAACAGTGA 176
Oy 1643 GGGTGAGTGTGACACCAAGAGACTGCTTCTGCTGCTTCCCTCAGCTCTCTTCAAGC 1702
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 175 GGGTGAGTGTGACACCAAGAGACTGCTTCAACTGCTTCACTGCTTCCCAACMAACRAC 116
Oy 1703 TTGGAGCTGACAT--GACGAGCCCATGATGCTGTGACAGAGTCCCTGCTC---TGTGGC 1757
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 115 TTGAAATCATMATAGACACRCCAMMGATCTGTGACAGAGTCTCTCTGATAGACACA 56
Oy 1758 AGGCTCTGTTTATAGC--CATGATCAGATGTGGTACAGACTTTTGTGGGCTTGA 1811
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 55 GGAGTCTCTTATAGACCAAGATCAGATGTGTGACAGCTCTTGTGGGCTTGA 1

```

```

RESULT 7
BQ446539/c 777 bp mRNA linear EST 29-MAY-2002
LOCUS BQ446539
DEFINITION UI-H-EU1-azx-9-10-0-UI-s1 NCI CGAP Ctl Homo sapiens cDNA clone
          UI-H-EU1-azx-9-10-0-UI 3', mRNA sequence.
ACCESSION BQ446539
VERSION BQ446539.1 GI:21249651
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 777)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
        Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
        Email: cgapds-tr@mail.nih.gov

```

Tissue Procurement: Dr. Jose Mercuende  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Sequencing by: Dr. M. Bento Soares, University of Iowa  
 The following repetitive elements were found in this cDNA  
 sequence: 1-50, >POLY\_A\$Simple\_repeat (matched complement)  
 Seq primer: M13 FORWARD  
 POLY\_A=Yes.

#### FEATURES

```

source location/Qualifiers
1..777
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-H-EU1-azx-9-10-0-UI"
/clone_lib="NCI CGAP Ctl"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Knee; Vector: p773-Pac (Pharmacia) with a
modified polylinker; Site 1: Ecor I; Site 2: Not I;
NCI CGAP Ctl is a normalized cDNA library containing the
following tissue(s): Osteoarthritic Cartilage The library
was constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into p773-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
(sequence that is located between the Not I site and the
(drr)18 tail). The sequence tag for this library is
TAGTCACGCT.
TAG_Lib=UI-H-EU1
TAG_TISSUE=osteothritic cartilage
TAG_SEQ=TAGTCACGCT"
BASE COUNT 168 a 208 c 194 g 203 t 4 others
ORIGIN

```

```

Query Match 33.5%; Score 672.6; DB 14; Length 777;
Best Local Similarity 96.9%; Pred No. 2.2e-112;
Matches 726; Conservative 0; Mismatches 18; Indels 5; Gaps 4;

```

```

Oy 1263 TTTTGTCCGAATGCCATCTAGGCGCAAGACCCCAAGGCGCAGAGTCTTC 1322
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 776 TTTTGTCCGAATGCCATCTAGGCGCAAGACCCCAAGGCGCAGAGTCTTC 717
Oy 1323 TTGCAGCAAAAGATGTGTGAAGACAGGCTTGCATGCGTGGTGGCCGCAAAATTCAT 1382
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 716 TTGCAGCAAAAGATGTGTGAAGACAGGCTTGCATGCGTGGTGGCCGCAAAATTCAT 657
Oy 1383 CCAGCTTACTGACACAGTATGATCGCCCATTTGCTGTTAAGCCCAAGAGTTTA 1442
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 656 CCAGCTTACTGACACAGTATGATCGCCCATTTGCTGTTAAGCCCAAGAGTTTA 597
Oy 1443 GCAGAGGAGTGAAGGAGAAACCGTCTCTCTTCCAGGCCAAGCTGACACTTCT 1502
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 596 GCAGAGGAGTGAAGGAGAAACCGTCTCTCTTCCAGGCCAAGCTGACACTTCT 537
Oy 1503 GTGGTGTGAGAAACAGTGTGACCAATGGCCAGACACATGCTTGCACATTTGAAGTC 1562
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 536 GTGGTGTGAGAAACAGTGTGACCAATGGCCAGACACATGCTTGCACATTTGAAGTC 477
Oy 1563 AAGGACTGTCCAGGAGAGG--TTTGTGACACAGGCTTCTTGGGAGCAGCTGAGGT 1620
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 476 AAGGACTGTCCAGGAGAGGTTTGTGACACAGGCTTCTTGGGAGCAGCTGAGGT 417
Oy 1621 GTAAAGGAGACAAACAGGTGAGGTGAGTGTGACACCCAGAGAGTGTCTTCCGCCCT 1680
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 416 GTAAAGGAGACAAACAGGTGAGGTGAGTGTGACACCCAGAGAGTGTCTTCCGCCCT 357
Oy 1681 CACCTGCCCCCACTCTACGATGGAGGTGACATGACAGGCCCACTGATCTGTACAGA 1740
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```



```
Db 356 CACCCTGCCCCACTCCTACGACTGGAGAGTACATGACCCAGCCACTGATCTGTGACGA 297
Qy 1741 GGTCTGCT-CTGTGGCAGGCTCTGTTATATAGCCATGATGATGCTGACACTT 1799
Db 296 GGTCTGCTCTCTGTGGCAGGCTCTGTTATATAGCCATGATGATGCTGACACTT 237
Qy 1800 TCTGGGCTGGAGACACGCTCACTTGTGACTGTCTGTGGACAGAGATGCTTGAGGC 1859
Db 236 TCTGGGCTGGAGACACGCTCACTTGTGACTGTCTGTGGACAGAGATGCTTGAGGC 177
Qy 1860 ATTCAGGACGCTCAGACCCAGCTTCTACCTGCTTGTGACTTCTCTA-GCATAGCCT 1918
Db 176 ATTCAGGACGCTCAGACCCAGCTTCTACCTGCTTGTGACTTCTCTAAGCATAGCCT 117
Qy 1919 GGGCAGGAGGCTGGGAGATGAGATGACATGGAGATGATGAGAGATGATGAT 1978
Db 116 GGGCAGGAGGCTGGGAGATGAGATGAGATGAGATGAGATGAGATGATGAT 58
Qy 1979 TTCCCGAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 2007
Db 57 TTCTGTTAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 29

RESULT 8
BF038722 769 bp mRNA linear EST 20-OCT-2000
LOCUS 601460641F1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3864082 5',
DEFINITION mRNA sequence.
ACCESSION BF038722 GI:10745770
VERSION BF038722.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 769)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9605 row: a column: 11
High quality sequence stop: 632.
Location/Qualifiers
1.769
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3864082"
/clone.lib="NIH_MGC_66"
/tissue.type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: ovary; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
BASE COUNT 184 a 212 c 217 g 156 t
ORIGIN
Query Match 32.9%; Score 659.8; DB 12; Length 769;
Best Local Similarity 97.8%; Pred. No. 4.7e-110;
Matches 701; Conservative 0; Mismatches 12; Indels 4; Gaps 3;
```

```
Qy 930 AACCTGAACCTCCACCACTGAGAGTGTGAGATGCTGCTTTTAAAGAGCCC 989
Db 1 AACCTGAACCTCCACCACTGAGAGTGTGAGATGCTGCTTTTAAAGAGCCC 60
```

```
Qy 990 AGCCCTGTGAGAGTGAATCTGAAGCTCCGCCGACCTTCCTCGATATATGATCT 1049
Db 61 AGCCCTGTGAGAGTGAATCTGAAGCTCCGCCGACCTTCCTCGATATATGATCT 120
Qy 1050 CAATGCTACCTTTGATGTGATATCTCCAGCCGAGCCCTCCAGCTCCAGATGGTTA 1109
Db 121 CAATGCTACCTTTGATGTGATATCTCCAGCCGAGCCCTCCAGCTCCAGATGGTTA 180
Qy 1110 CTACGAAAAAATTTGCTTAAGAAATGACACACTCCCAATTCAGATGTCCCAAGAAGAT 1169
Db 181 CTACGAAAAAATTTGCTTAAGAAATGACACACTCCCAATTCAGATGTCCCAAGAAGAT 240
Qy 1170 ATCAAAAGGCCAGAGAGATGCCAGCTCTCACTGAGTGGCCAGAGCTGTGAGAGA 1229
Db 241 ATCAAAAGGCCAGAGAGATGCCAGCTCTCACTGAGTGGCCAGAGCTGTGAGAGA 300
Qy 1230 GCCAGATGAGAACTGGTGGCTCCCTATTTTGTCCGGAATGCCATCTAGGCCA 1289
Db 301 GCCAGATGAGAACTGGTGGCTCCCTATTTTGTCCGGAATGCCATCTAGGCCA 360
Qy 1290 GAAACAGCCCAAAAGGCCAGAGTCAAGTCTCTGACAGAAAGATGTGTAAGACAG 1349
Db 361 GAAACAGCCCAAAAGGCCAGAGTCAAGTCTCTGACAGAAAGATGTGTAAGACAG 420
Qy 1350 CTTCGATGGCTGGTGGCCGACAAATTCATCCAGCTTCACTGACAGATCATGATCG 1409
Db 421 CTTCGATGGCTGGTGGCCGACAAATTCATCCAGCTTCACTGACAGATCATGATCG 480
Qy 1410 CCATTTGCTTTAAGCCCAAGACCAAGTTAAGCAGAGGATGAGGTTAAGACCTGCC 1469
Db 481 CCATTTGCTTTAAGCCCAAGACCAAGTTAAGCAGAGGATGAGGTTAAGACCTGCC 540
Qy 1470 TTCTCTCTCCAGCCCAAGCTGACACCTTCTGTGTCTGTGAAACAGTGTGACC 1529
Db 541 TTCTCTCTCCAGCCCAAGCTGACACCTTCTGTGTCTGTGAAACAGTGTGACC 600
Qy 1530 AATGGCAGACATGCTCACTTGTAGTGAAGAGCTGTCAGAGAGG--TTGTG 1587
Db 601 AATGGCAGACATGCTCACTTCAAC-TCATGCTTAAGAGCTGTCCAGAGGCTTTGTG 659
Qy 1588 GACAGACCCCTACTTTCGGGAGCAGCCTGAGGTGAAGCAACAAGAGTGAAG 1644
Db 660 GACAGACCCCACTTTCGGGAGAG-CTGAGGTGTGTGGCAGACACAGTGCAGG 715

RESULT 9
BG764118 702 bp mRNA linear EST 15-MAY-2001
LOCUS 602737137F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4862312 5',
DEFINITION mRNA sequence.
ACCESSION BG764118 GI:14074771
VERSION BG764118
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 702)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM1722 row: j column: 09
High quality sequence stop: 697.
Location/Qualifiers
```

Source	1. 702
Query Match	32.6%; Score 654.2; DB 12; Length 702;
Best Local Similarity	97.1%; Pred. No. 5e-109;
Matches 676; Conservative	0; Mismatches 19; Indels 1; Gaps 1;
BASE COUNT	183 a 180 c 202 g 136 t 1 others
ORIGIN	Library.1
1. 702	<p> /organism="Homo sapiens"  /db_xref="taxon:9606"  /clone="IMAGE:4862312"  /clone_1lb="NIH-MGC_49"  /lssue.type="melanotic melanoma, high MDR (cell line)"  /lab_host="DH10B (phage-resistant)"  /Note="Organ: skin; Vector: pORF7; Site.1: XhoI; Site.2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCAGGAG(G). Size-selected &gt;500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library.1" </p>
33	GGGGGGCGGTACGACGCGGACCTGTAGACGTTCTTGGCTCGCGGGCCCTTAG 92
3	GGGGGGCGGTACGACGCGGACCTGTAGACGTTCTTGGCTCGCGGGCCCTTAG 62
93	TCCAGCCATCATGCTTATCCGCTCTGTGCACATATGCTCCGACTTTCGATCACTC 152
63	TCCAGCCATCATGCTTATCCGCTCTGTGCACATATGCTCCGACTTTCGATCACTC 122
133	CCGCGACGCTGGCCCGCATTCACCTTCGACACCTTCACATTCGATCACTC 212
123	CCGCGACGCTGGCCCGCATTCACCTTCGACACCTTCACATTCGATCACTC 182
213	CTTTCAGACGACCAAGTCGACCTGCGCCACAGTCCGGAATCCAGTTCGCAAAAGAC 272
183	GTTTCAGACGACCAAGTCGACCTGCGCCACAGTTCGCAAAAGAC 242
273	CATTTCATTAAGCTCTTCTTTGATCTTGCCAGAGAGAGAGAAATGCTTGCATCGACA 332
243	CATTTCATTAAGCTCTTCTTTGATCTTGCCAGAGAGAGAGAAATGCTTGCATCGACA 302
333	ATTTCATTAAGATGAAGTGCATATGTACAGACCCAGCTTTCGCAAAAGAGAGAA 392
303	ATTTCATTAAGATGAAGTGCATATGTACAGACCCAGCTTTCGCAAAAGAGAGAA 362
393	ACGACACAGCCAGGTCATCATGACACCTGCGGGATACGCTGGAAAGAACCAATGCTAC 452
363	ACGACACAGCCAGGTCATCATGACACCTGCGGGATACGCTGGAAAGAACCAATGCTAC 422
453	TGTGTATCTCTGCAGACGCTTGGGCAAGGCCAGATGCTGTCCTCCACACTGAAAAA 512
423	TGTGTATCTCTGCAGACGCTTGGGCAAGGCCAGATGCTGTCCTCCACACTGAAAAA 482
513	GCAGATGAAGTACTTAGAGCAGCAGCAGATGAGACCAAGACCAAGAGAGAGCGGG 572
483	GCAGATGAAGTACTTAGAGCAGCAGCAGATGAGACCAAGACCAAGAGAGAGCGGG 542
573	CCGCGTCAGAGCAAGATGAAGACCATGGAGCAGATTAGCTTCTACTCCAGAGCCACT 632
543	CCGCGTCAGAGCAAGATGAAGACCATGGAGCAGATTAGCTTCTACTCCAGAGCCACT 602
633	CCCTCAGAGTGGAGGAG-GATGATCCAGACATGGAGTGGAGACATGACCGGTGGAACAGC 691
603	CCCTCAGAGTGGAGGAGCAGATGATCCAGACATGGAGTGGAGACATGACCGGTGGAACAGC 662
692	TGGCTGTACTGTGTCTCTCAAGAAAGATAG 727
663	TGGCTGTACTGTGTCTCTCAAGAAAGATAG 698

Accession	Definition	Accession	Definition	Accession	Definition
BE869186	BE869186	888 bp	mRNA	linear	EST 20-OCT-2000
LOCUS	601445180P1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3849451 5'				
DEFINITION	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
ACCESSION	BE869186				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)				
TITLE	Unpublished (1999)				
JOURNAL	Contact: Robert Strausberg, Ph.D.				
COMMENT	Email: cgabbs@email.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LML at: <a href="http://image.lml.gov">http://image.lml.gov</a> plate: LMA9566 row: 0 column: 20 High quality sequence stop: 692. Location/Qualifiers 1..888 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3849451" /clone_1id="NIH_MGC_65" /tissue_type="adenocarcinoma" /lab_host="DH10B (phage-resistant)" /note="Organ: colon; Vector: PCMV-SPOrt6; Site:1; Noit; Site-2: Salt; Cloned unidirectionally. Primer: oligo dt. Average insert size 1.8 kb. Library constructed by Life Technologies."				
BASE COUNT	254 a 199 c 268 g 167 t				
ORIGIN					
Query Match	32.5%	Score 651.4:	DB 12:	Length 888:	
Best Local Similarity	93.7%:	Pred. Mis:1.5e-108:			
Matches 788:	Conservative	0: Mismatches 41:	Indels 12:	Gaps 10:	
OY	309 GGAGAGAATGCTGTGGATCGAATTCTTTAA-GAATGAAGTGGCAATGTCAGAGCCC	367			
Db	7 GGAGAGAAATGCTTGATCGAATTTCTTAAGGAATGAAGTGGCAATGTCAGAGCCC	66			
OY	368 AGCTTTCCAGAAAGACAAGAGAAAGAGACAGCAGCAGTGCATCATGACACTCTGCCGG	427			
Db	67 AGCTTTCCAGAAAGACAAGAGAAAGAGACAGCAGCAGTGCATCATGACACTCTGCCGG	126			
OY	428 ATACGCTGGAAGAACCC-AATGCTACTGTGGTATCTCTGACAGAGGCTTGGGCAAGGCC	486			
Db	127 ATACGCTGGAAGAAAGCAATGCTACTGTGGTATCTCTGACAGAGGCTTGGGCAAGGCC	186			
OY	487 GAGATGCTGTGCTCCACTGAAAAAGAGAGATGAAGTACTTAGAGCAGCAGCAGAGATGAG	546			
Db	187 GAGATGCTGTGCTCCACTGAAAAAGAGAGATGAAGTACTTAGAGCAGCAGCAGAGATGAG	246			
OY	547 ACCAAACAAGACAAGAGAGAGGCGCGGCTCAGAGACAAGATGAAGACCATGGAGCAG	606			
Db	247 ACCAAACAAGACAAGAGAGAGGCGCGGCTCAGAGACAAGATGAAGACCATGGAGCAG	306			
OY	607 ATTGAGCTTCTACTCCAAGCCAGCTCCCTGAGGTGAGGAGATGATCCGAGACATGGGT	666			
Db	307 ATTGAGCTTCTACTCCAAGCCAGGCGCTGAGGTGAGGAGATGATCCGAGACATGGGT	366			
OY	667 GTGGACAGTCAAGCGGTGGAACAGCTGGCTGTGTAATCTGTGCTCTCAAAAAGAGATC	726			
Db	367 GTGGACAGTCAAGCGGTGGAACAGCTGGCTGTGTAATCTGTGCTCTCAAAAAGAGATC	426			

ORIGIN	BASE COUNT	207 a	200 c	224 g	153 t	
Query Match	32.4%					Score 651; DB 12; Length 784;
Best Local Similarity	93.6%					Pred. No. 1.8e-108;
Matches 735; Conservative	0;					Mismatches 40; Indels 10; Gaps 5.
Qy	16	ATTTAACCAAGCGGAGGC-----GGGGCGCTCTACGAAGCCGGAGCTGTACAGTTTCT	70			
Db	1	ATTTAAACCAACGGAGGCGCGCGCGCGCTGTACGAAGCCGGAGCTGTACAGTTTCT	60			
Qy	71	TTGGCTGCGTGGGGCCCTTTGAGTCCAGGCATCATGCTATCCGTCCTGTGCACATATCT	130			
Db	61	TTGGCTGCGTGGGGCCCTTTGAGTCCAGGCATCATGCTATCCGTCCTGTGCACATATCT	120			
Qy	131	GCTCCGACTTCTTGCATCACTCCGCGAGCTGGCGGCATCCACTATCGGCGCACACCTTCC	190			
Db	121	GCTCCGACTTCTTGCATCACTCCGCGAGCTGGCGGCATCCACTATCGGCGCACACCTTCC	180			
Qy	191	ACTTCCAGGCGCTAATTACATGCTCTGAGACGACACCAAGTGGGACCTGCCACATGTCC	250			
Db	181	ACTTCCAGGCGCTAATTACATGCTCTGAGACGACACCAAGTGGGACCTGCCACATGTCC	240			
Qy	251	GAATCCAGTTGGCAAAAGAACCATTTATCAATAAGCTCTTCTTGAATCTTGGCCAGGAG	310			
Db	241	GAATCCAGTTGGCAAAAGAACCATTTATCAATAAGCTCTTCTTGAATCTTGGCCAGGAG	300			
Qy	311	AGGAGAATGTCTTGGATCGAGAAATTTCTTAAGAATGAACTGGACAATGTCCAGAGCCACG	370			
Db	301	AGGAGAATGTCTTGGATCGAGAAATTTCTTAAGAATGAACTGGACAATGTCCAGAGCCACG	360			

Db	361	TTTCCAGAAACACAGCAAGAAACGAGACGCCAGTCAATCATCCACACTCTGCGGATA	420
OY	421	CGCTGAGAAACGCAATGCTACTGTGTGTAATCTCTGACGACGAGCCTTTGGCAAGCCGAGA	490
Db	421	CGCTGGAGAAACGCATGCTACTGTGTGTAATCTCTGACGACGAGCCTTTGGCAAGCCGAGA	480
OY	491	TGCTGTGCTCCCACTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGCAGATGAGACCA	550
Db	481	TGCTGTGCTCCCACTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGCAGATGAGACCA	540
OY	551	AACAGACACAAAGAGAGGGGGGCGGCTCAGAGCAAGATGAAGACATGAGACAGATTG	610
Db	541	AACAGACACAAAGAGAGGGGCCG -CGGCTTAGAGACAAAGATGAAGACATGAGACAGATTG	599
OY	611	AGCTTCTACTCCAGAGCCAGCTCCCTGAGAGTGGAGAGATGATCCGAGACATGGGTGTG	670
Db	600	AGCTTCTACTCCAGAGCCAGCCCTCGAGAGTGGAGAGATGATCCGAGACATGGGTGTG	659
OY	671	GACACTCAGCGGTGGAACAGCTGCGTGTGTACTGTGTCTCTCAAGAAAGAGTACGACA	730
Db	660	GACACTCAGCGGTGGAACCGCTGG -TGTTACTGTGTGTCTCTCCGAAAGAGT -CGAAA	717
OY	731	ATCTAAAGAGCAGGAGGACCTCAGAGGAGGTGGCTCAACAGCTGAGGAAGAGATTGT	790
Db	718	ATCTAAAGAGCAGGAGGACCTCAGAGGAGGTGGCTCAACAGCTGAGGAAGAGATTGT	775
OY	791	TTTTC 795'	
Db	776	CCTCC 780	

RESULT_12					
BC682548	BC682548	752 bp	mRNA	linear	EST: 01-MAY-2001
LOCUS	6026224562E1	NCI_CGAP_Skn4	Homo sapiens	CDNA clone	IMAGE:4749447 5'
DEFINITION	mRNA sequence.				
ACCESSION	BC682548				
VERSION	BC682548.1	GI:13913945			
KEYWORDS	EST.				
SOURCE	human.				

```

ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE          NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL        National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT        Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-remail.nih.gov
                Tissue Procurement: James Cleaver, M.D.
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLM10602 row: k column: 16
                High quality sequence start: 19
                High quality sequence stop: 736.
                Location/Qualifiers
                1..752
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_image="4749447"
                /clone_lib="NCI_CGAP_Skn4"
                /tissue_type="squamous cell carcinoma"
                /lab_host="DH10B (T1 phage-resistant)"
                /note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
                Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.5kb. Library constructed by Life
                Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT    200 a 194 c 212 g 146 t
ORIGIN
Query Match      32.4% Score 650.2; DB 12; Length 752;
Best Local Similarity 97.0%; Pred. No. 2.6e-108;
Matches 705; Conservative 0; Mismatches 18; Indels 4; Gaps 4;

37 GGGGCTACGAGAGCGGAGCTGTAGCTTTTGGTGGCTGCTGGCCCTTGAGTCCA 96
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
22 GGGCGCTGAGAGCGGAGCTGTAGCTTTTGGTGGCTGCTGGCCCTTGAGTCCA 80
97 GCCATATGCTATCGCTGCTGTGCACTATCTGCTCGACTTCTTGATCAGTCCCGC 156
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
81 GCCATATGCTATCGCTGCTGTGCACTATCTGCTCGACTTCTTGATCAGTCCCGC 140
157 GAGTGGCGGCGCATGCACTGCGGCGACACTTGCATTTGAGTGCCTAATTCAGTCTT 216
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
141 GAGTGGCGGCGCATGCACTGCGGCGACACTTGCATTTGAGTGCCTAATTCAGTCTT 200
217 GAGACAGCACAAGTGGAGCTGCCAGAGTCCGATCCAGG-TTGGCAAAAGAACAT 275
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
201 GAGACAGCACAAGTGGAGCTGCCAGAGTCCGATCCAGGTTTGGCAAAAGAACAT 260
276 TATCAATTAAGCTCTTTTGATCTTCCAGAGAGAGAGATGCTTGGATCGAAT 335
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
261 TATCAATTAAGCTCTTTTGATCTTCCAGAGAGAGAGATGCTTGGATCGAAT 320
336 CTTAAGANTGAAGTGGACAATGTGAGAGCCGCTTCCCAAGAGACAGAGAAAGC 395
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
321 CTTAAGANTGAAGTGGACAATGTGAGAGCCCGCTTCCCAAGAGACAGAGAAAGC 380
396 AGACAGCACAAGTGGAGCTGCCAGAGTCCGATCCAGG-TTGGCAAAAGAACAT 455
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
381 AGACAGCACAAGTGGAGCTGCCAGAGTCCGATCCAGGTTTGGCAAAAGAACAT 440
456 GGTATCTTGACAGCAGGCTTGAGGCAAGCGGAGATGCTGTGCTCCACTGAAAAGCA 515
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
441 GGTATCTTGACAGCAGGCTTGAGGCAAGCGGAGATGCTGTGCTCCACTGAAAAGCA 500
516 GATGAGTACTTATGAGCAGCAGAGATGAGACCAAGCAGCAGAGAGAGAGCGCGG 575
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
501 GATGAGTACTTATGAGCAGCAGAGATGAGACCAAGCAGCAGAGAGAGAGCGCGG 560

```

```

OY 576 GCTCAGACCAAGATGAGACCATGAGCAGATTGAGCTTCTACTCCAGAGCTCCC 635
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 561 GCTCAGACCAAGATGAGACCATGAGCAGATTGAGCTTCTACTCCAGAGCAGCGCC 620
OY 636 TGAGGTGAGAGATGATCCGAGACATGGGTGTGGACAGTACGCGTGG -AACACTGG 694
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 621 TGAGGTGAGAGATGATCCGAGACATGGGTGTGGACAGCAGCGGTGAAAAGTGG 680
OY 695 CTGTGACGTGTGTCTC-CAAGAAAGATCGAGAACTTAAGAGGCGAGAAAGGCC 753
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 681 CTGTGACGTGTGTCTC-CAAGAAAGATCGAGAACTTAAGAGGCGAGAAAGGCC 740
OY 754 TGAGGCG 760
11 | | | | |
DB 741 TCCGGCG 747

RESULT 13
BG420765      908 bp mRNA linear EST 14-MAR-2001
LOCUS        602448558p1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4586846 5',
DEFINITION   mRNA sequence.
ACCESSION    BG420765
VERSION      BG420765.1 GI:13327271
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE        NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-remail.nih.gov
                Tissue Procurement: DCTD/drp
                cDNA Library Preparation: Ling Hong/Rubin Laboratory
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LCM1316 row: d column: 15
                High quality sequence stop: 719.
                Location/Qualifiers
                1..908
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_image="4586846"
                /clone_lib="NIH_MGC_14"
                /tissue_type="renal cell adenocarcinoma"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: kidney; Vector: pOTB1; Site:1: XhoI; Site:2:
                EcoRI; cDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GCCACGAG(G). Size-selected >500bp for average
                insert size 1.8kb. Library constructed by Ling Hong in
                the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT    235 a 229 c 274 g 170 t
ORIGIN
Query Match      31.5% Score 632.6; DB 12; Length 908;
Best Local Similarity 89.2%; Pred. No. 3.9e-105;
Matches 770; Conservative 0; Mismatches 39; Indels 54; Gaps 6;

82 GGGCCCTTATGTCACCACTATGCTATCCGCTGCTGTGACTATGCTCCGACTTC 141
11 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 GGGCCCTTATGTCACCACTATGCTATCCGCTGCTGTGACTATGCTCCGACTTC 60
OY 142 TTGATCACTCCCGGAGCTGGCGGCATCACTGGCGCACACCTTCCACTTGGAGTGC 201

```



```

QY 608 TTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGAGAGATGATCCGAGACATGGGTG 667
DB 601 TTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGAGAGATGATCCGAGACATGGGTG 660
QY 668 TGGGACAG 675
DB 661 TGGGACAG 668

RESULT 15
AL559248 645 bp mRNA linear EST 16-FEB-2001
LOCUS AL559248 LTI_NFL008_TC2 Homo sapiens cDNA clone CS0DJ012YH21 5
DEFINITION prime, mRNA sequence.
ACCESSION AL559248
VERSION AL559248.1 GI:12904563
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 645)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 645
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DJ012YH21"
/clone_1ib="LTI_NFL008_TC2"
/sex="male"
/tissue_type="T cells from T cell leukemia"
/note="Vector: pCMVSPORT 6; site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 167 a 174 c 177 g 127 t
ORIGIN

Query Match 31.4%; Score 629.8; DB 9; Length 645;
Best Local Similarity 98.9%; Pred. No. 1.4e-104;
Matches 634; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 42 TCTACGAAGCCGAGCTGTAGACGTTCTTGGCTGCTGGGCCCTTGAGTCAGCCAT 101
DB 5 TCTACGAAGCCGAGCTGTAGACGTTCTTGGCTGCTGGGCCCTTGAGTCAGCCAT 64
QY 102 CATGCTATCCGCTGCTGTGCTATGCTGCTCCGACTTCTTGATCACTCCCGCAGCT 161
DB 65 CATGCTATCCGCTGCTGTGCTATGCTGCTCCGACTTCTTGATCACTCCCGCAGCT 124
QY 162 GGGCGCCATCCATGCGGCGACACTTCCACTTGAGTGCGCTAATTGAGCTTTGAGAC 221
DB 125 GGGCGCCATCCATGCGGCGACACTTCCACTTGAGTGCGCTAATTGAGCTTTGAGAC 184
QY 222 AGCAGCAAGTCGAGCTGCCACAGTGCAGATCCAGGTTGGCAAAAGAACATTATCAA 281
DB 185 AGCAGCAAGTCGAGCTGCCACAGTGCAGATCCAGGTTGGCAAAAGAACATTATCAA 244
QY 282 TAAAGCTCTTCTTGATCTTGCCAGAGAGAGAAATGCTTGATGAGAAATTTTAA 341
DB 245 TAAAGCTCTTCTTGATCTTGCCAGAGAGAGAAATGCTTGATGAGAAATTTTAA 304

```

```

QY 342 GAATGAAGTGAACATGTGAGAGCCAGCTTCCAGAAAAGCAAGAGAAAGAGACAG 401
DB 305 GAATGAAGTGAACATGTGAGAGCCAGCTTCCAGAAAAGCAAGAGAAAGAGACAG 364
QY 402 CCAGGTCAATCATGAGACACTCTGCGGAGTACGCTGAAGAAGCAATGCTACTGTGATC 461
DB 365 CCAGGTCAATCATGAGACACTCTGCGGAGTACGCTGAAGAAGCAATGCTACTGTGATC 424
QY 462 TCTGCAAGAGGCTTGGGCAAGGCCAGATGCTGTCTCCACACTGAAAAAGCAGATGAA 521
DB 425 TCTGCAAGAGGCTTGGGCAAGGCCAGATGCTGTCTCCACACTGAAAAAGCAGATGAA 484
QY 522 GTACTTAGAGACAGCAGAGATGAGACAAACCAAGCAGAGAGAGCGCGCGCTCAG 581
DB 485 GTACTTAGAGACAGCAGAGATGAGACAAACCAAGCAGAGAGAGCGCGCGCTCAG 544
QY 582 GAGCAAGATGAAGACCATGAGAGCAATGAGCTTCTACTCCAGAGCCAGCTCCAGGT 641
DB 545 GAGCAAGATGAAGACCATGAGAGCAATGAGCTTCTACTCCAGAGCCAGCTCCAGGT 604
QY 642 GGAGGAGATGATCCGAGACATGGGTGGGACAGTCAGCGG 682
DB 605 GGAGGAGATGATCCGAGACATGGGTGGGACAGTCAGCGG 645

```

Search completed: December 13, 2002, 04:37:28  
Job time : 1982 secs